

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:26:02 ; Search time 29.37 Seconds

(without alignments)
31.887 Million cell updates/sec

Title: US-09-331-631A-33

Perfect score: 77
Sequence: 1 CXXCXXCXKXXXXXXXXXXCXKXXCXKXXC 29

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77	100.0	43	1	MT2_DROME
2	68	88.3	55	1	M84C_DROME
3	68	88.3	63	1	M84D_DROME
4	68	88.3	68	1	M84D_DROME
5	68	88.3	74	1	M84B_DROME
6	68	88.3	74	1	MDNM_MOUSE
7	68	88.3	74	1	MDNM_MOUSE
8	68	88.3	95	1	EP4_RABIT
9	68	88.3	96	1	EP4_RABIT
10	68	88.3	99	1	RS11_LIGES
11	68	88.3	98	1	GAS1_ARATH
12	68	88.3	99	1	GAS2_ARATH
13	68	88.3	106	1	GAS4_ARATH
14	68	88.3	110	1	IBP_CARCR
15	68	88.3	112	1	IBP_CARCR
16	68	88.3	124	1	EP4_CANFA
17	68	88.3	125	1	EP4_HUMAN
18	68	88.3	131	1	VWF_BOVIN
19	68	88.3	157	1	VES_RHEV1
20	68	88.3	169	1	KRUA_HUMAN
21	68	88.3	182	1	KRUC_SHEEP
22	68	88.3	210	1	YOEH_ECOLI
23	68	88.3	676	1	KALM_CHICK
24	68	88.3	680	1	KALM_CHICK
25	68	88.3	703	1	EGFR_CHICK
26	68	88.3	1166	1	XMRC_XIPMA
27	68	88.3	1210	1	EGFR_HUMAN
28	68	88.3	1210	1	EGFR_MOUSE
29	68	88.3	1308	1	ERB4_HUMAN
30	68	88.3	1339	1	ERB3_RAT
31	68	88.3	1342	1	ERB3_RAT
32	68	88.3	1700	1	BAR3_HUMAN
33	68	88.3	2476	1	ZAN_PIG

34	68	88.3	2813	1	VWF_CANFA	Q28295 canis fam11
35	68	88.3	7813	1	VWF_HUMAN	P04275 homo sapien
36	67	87.0	71	1	MT21_MYTED	P80251 mytilus edu
37	67	87.0	71	1	MT22_MYTED	P80252 mytilus edu
38	67	87.0	71	1	MT23_MYTED	P80253 mytilus edu
39	67	87.0	642	1	YB55_SCHPO	Q09746 schizosacch
40	67	87.0	704	1	KPC1_CAEL	P34722 caenorhabd1
41	67	87.0	912	1	KPCM_HUMAN	Q15139 homo sapien
42	67	87.0	988	1	PKC1_SCHPO	P36582 schizosacch
43	67	87.0	1096	1	KPC1_ASPNG	Q00078 aspergillus
44	67	87.0	1097	1	KPC1_CANAL	P43057 canida alb
45	67	87.0	1142	1	VGLM_PHV	P27315 prospect hi

ALIGNMENTS

RESULT	ID	MT2_DROME	STANDARD	PRT	43 AA.
AC	P11956	Q9VDM2:			
DT	01-OCT-1989	(Rel. 12, Created)			
DT	01-OCT-1989	(Rel. 12, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	METALLOTHIONEIN 2 (MT-2).				
GN	MTNB OR MTO.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=OREGON-R;				
RX	MEDLINE=87204190; PubMed=3106973;				
RA	Mokdad R., Debec A., Wegnez M.;				
RT	"Metallothionein genes in Drosophila melanogaster constitute a dual				
RT	system.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 84:2658-2662(1987).				
RN	(2)				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 3-27.				
RC	STRAIN=OREGON-R;				
RX	MEDLINE=91012582; PubMed=1976815;				
RA	Silar P., Theodore L., Mokdad R., Erraiss N.-E., Cadic A., Wegnez M.;				
RT	"Metallothionein Mto gene of Drosophila melanogaster: structure and				
RT	regulation.";				
RL	J. Mol. Biol. 215:217-224(1990).				
RN	(3)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BERKELEY;				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Miller B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., McKelvey G.L.G.,				
RA	Abrell J.F., Abmayyan A., An H.-J., Andrews-Pfankuch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borrows D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,				
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,				
RA	Foster C., Gaitanellis A.E., Gang N.S., Gelbart W.M., Glasser K.,				
RA	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harits N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,				
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasto P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				

```

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
CC Science 287:2185-2195(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS CATIONS OF SEVERAL TRANSITION
CC ELEMENTS. THOUGHT TO BE INVOLVED IN METAL ION HOMEOSTASIS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED PREDOMINANTLY IN EMBRYONIC AND
CC LARVAL STAGES.
CC -1- INDUCTION: STRONGLY INDUCED BY CADMIUM, COPPER AND MERCURY.
CC -1- MISCELLANEOUS: ALL CYSTEINE RESIDUES ARE ARRANGED IN C-X-C GROUPS.
CC THESE ARE THOUGHT TO BE THE METAL-BINDING SITES IN OTHER
CC METALLOTHIONEINS.
CC -1- SIMILARITY: THE TWO DROSOPHILA METALLOTHIONEINS (MTO AND MTN) ARE
CC ONLY 26% SIMILAR.
CC -1- SIMILARITY: BELONGS TO FAMILY 5 IN METALLOTHIONEIN SUPERFAMILY.
CC -1- SIMILARITY: BELONGS TO FAMILY 5 IN METALLOTHIONEIN SUPERFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: M16250; AAA28683.1; -
DR EMBL: X52098; CA36318.1; -
DR EMBL: AE003730; AAF5768.1; -
DR PIR: S14706; SMF2.
DR FLYBASE: FBgn0002869; Mttn.
DR INTERPRO: IPR000966; -
DR PFM: PF02067; Metallothio_5; 1.
DR PRINTS: PR00872; MTNIPTRRA.
DR METAL-BINDING: Chelation; Copper; Cadmium; Zinc.
KW SEQUENCE 43 AA; 4525 MW; 5EE0CF9171BD9A97 CRC64;
SQ
Query Match 100.0%; Score 77; DB 1; Length 43;
Best Local Similarity 20.7%; Pred. No. 4.2;
Matches 6; Conservative 23; Mismatches 0; Indels 0; Gaps 0;
1 CXXCXXCXXCXXCXXCXXCXXCXXC 29
1 |:::|:::|:::|:::|:::|:::|
3 CKGCGTNCOCNAKCGDNACNKCQVC 31

```

```

RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
RA Schaefer M.;
RT "A cluster of four genes selectively expressed in the male germ line
RT of Drosophila melanogaster.";
CC Mech. Dev. 35:143-151(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck H., Brokstein P., Brothier P.,
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
CC Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
CC -1- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
CC MOTIFS.
CC -1- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC -1- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: X67703; CAA47939.1; -
DR EMBL: AE003672; AAF54025.1; -
DR HSSP: P01180; INPO.
DR FLYBASE: FBgn0004174; Mst84DC.
KW Spermatogenesis; Repeat; Multigene family.
KW SEQUENCE 55 AA; 5225 MW; 95A12F3AEC8BDB6C CRC64;
SQ
Query Match 88.3%; Score 68; DB 1; Length 55;
Best Local Similarity 17.9%; Pred. No. 25;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 CXXCXXXXXXXXXXXXXXXXXX 28
DB 19 CGCGPCGCGCGCGCGCGCGCG 46

RESULT 3
M84D_DROME STANDARD; PRT; 63 AA.
ID M84D_DROME
AC 001642; Q9VIA3;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MALE SPECIFIC SPERM PROTEIN MST84DA.
GN MST84DA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R;
RX MEDLINE=92102953; PubMed=1684716;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Arroll J.F., Agayari A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballif J.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,
RA Klamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Misha N.V., Mody B., Murphy L., Munz D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paczleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski R., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Stroup R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
CC -1- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
CC MOTIFS.

```

```

CC -1- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67703; CAA47937.1; -
DR EMBL: AE003672; AAF54022.1; -
DR HSSP: P01180; INPO.
DR FLYBASE: FBgn0004172; Mst84da.
KW Spermatogenesis; Repeat; Multigene family.
SQ SEQUENCE 63 AA; 5806 MW; BF84CD74CBCEFD3F CRC64;

Query Match 88.3%; Score 68; DB 1; Length 63;
Best Local Similarity 17.9%; Pred. No. 27;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

OY 1 CXXCXXXXXXXXXXXXXXXXXX 28
DB 26 CGCGPCGCGCGCGCGCGCGCG 53

RESULT 4
M84D_DROME STANDARD; PRT; 68 AA.
ID M84D_DROME
AC 001645;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MALE SPECIFIC SPERM PROTEIN MST84DD.
GN MST84DD.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OREGON-R;
RX MEDLINE=92102953; PubMed=1684716;
RA Kuhn R., Kuhn C., Boersch D., Glaetzer K.H., Schaefer U.,
RA Schaefer M.;
RT "A cluster of four genes selectively expressed in the male germ line
RT of Drosophila melanogaster.";
RL Mech. Dev. 35:143-151(1991).
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: PRIMARY SPERMATOCYTES.
CC -1- DOMAIN: THIS PROTEIN IS MOSTLY COMPOSED OF REPETITIVE C-G-P
CC MOTIFS.
CC -----
CC -1- SIMILARITY: BELONGS TO THE MST(3)CGP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67703; CAA47940.1; -
DR HSSP: P01180; INPO.
DR FLYBASE: FBgn0004175; Mst84dd.
KW Spermatogenesis; Repeat; Multigene family.
SQ SEQUENCE 68 AA; 6480 MW; 2F2BD26128DE3DDE CRC64;

Query Match 88.3%; Score 68; DB 1; Length 68;
Best Local Similarity 17.9%; Pred. No. 29;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]

```

RA MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "nec and ras initiate murine mammary tumors that share genetic
RT markers generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
CC -I- FUNCTION: COULD HAVE PROTEINASE INHIBITING CAPACITY.
CC -I- SIMILARITY: BELONGS TO THE MAP-TYPE 'FOUR-DISULFIDE CORE' FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL: X93037; CAA63605.1; -.
DR HSSP: P19957; 2REL.
DR MGD: MG1:107506; EXPL.
DR INTERPRO: IPR002221; -.
DR PFAM: PF00095; wap; 1.
DR PRINTS: PR00003; 4DISULPHCORE.
DR PROSITE: PS00317; 4_DISULFIDE_CORE; 1.
KW Protease inhibitor; Signal.
FT SIGNAL 1 24
FT CHAIN 25 74 POTENTIAL.
FT SEQUENCE 74 AA; 7787 MW; 42BDCAD5B330B3D3 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 74;
Best Local Similarity 17.9%; Pred No. 31;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0.

QY 2 XXXXXXXXXXXXXXXXXXXCCXXCXXC 29
DB 42 GTCDERCTGDSGSCGNMCKCSNGCGHAC 69
.::::|::::::::::|:::|
.::::|::::::::::|:::|

RESULT 7
WDNM_RAT STANDARD: PRT; 74 AA.
ID WDNM_RAT PI4730;
AC 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE WDNM1 PROTEIN PRECURSOR (FRAGMENT).
GN WDNM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST CARCINOMA;
RX MEDLINE=88310901; PubMed=3136918;
RA Dear T.N., Kefford R.F.;
RA "The WDNM1 gene product is a novel member of the 'four-disulphide
RT core' family of proteins";
RT Blochem. Biophys. Res. Commun. 176:247-254(1991).
CC -I- FUNCTION: INVOLVED IN THE METASTATIC POTENTIAL OF ADENOCARCINOMAS
CC IN RAT. COULD HAVE PROTEINASE INHIBITING CAPACITY.
CC -I- TISSUE SPECIFICITY: SPLEEN, KIDNEY, BRAIN, LIVER, LUNG, AND CELL
CC LINE RAT-2.
CC -I- SIMILARITY: BELONGS TO THE MAP-TYPE 'FOUR-DISULFIDE CORE' FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X13309; CA31688.1; ALT_INIT.
 DR PIR: JH0390; JH0390.
 DR HSSP: P19957; 2REL.
 DR INTERPRO: IPR002221; -.
 DR PFAM: PF00095; Wap: 1.
 DR PROSITE: PS00317; 4_DISULFIDE_CORE; 1.
 KW Protease inhibitor; Signal.
 FT NON_TER 1 1
 FT SIGNAL 23
 FT CHAIN 24 74 MDNM1 PROTEIN.
 SO SEQUENCE 74 AA; 7740 MW; 51802C70CDADF0521 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 74;
 Best Local Similarity 17.9%; Pred. No. 31;
 Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

OY 2 XXXXXXXXXXXXXXXXXXXXXXXC 29
 DB 42 GTCVELCSGDQSCPNIQKCCNCGHVC 69

RESULT 8
 ID EP4_RABIT STANDARD; PRT; 95 AA.
 AC 028631;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE MAJOR EPIDIDYMIS-SPECIFIC PROTEIN E4 (EPIDIDYMAL PROTEIN BE-20)
 DE (FRAMEMENT).
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EPIDIDYMIS;
 RA Wang L.F.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: POSSIBLE FUNCTION IN SPERM MATURATION.
 CC -1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: EPIDIDYMIS.
 CC -1- SIMILARITY: BELONGS TO THE WAP-TYPE 'FOUR-DISULFIDE CORE' FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U26725; AAA6525.1; -
 DR HSSP: P19957; 1ELE.
 DR INTERPRO: IPR002221; -.
 DR PFAM: PF00095; Wap: 2.
 DR PROSITE: PS00317; 4_DISULFIDE_CORE; 2.
 KW Glycoprotein.
 FT NON_TER 1 1
 FT DOMAIN <1 47 WAP 1.
 FT DOMAIN 48 95 WAP 2.
 FT CARBOHYD 17 17 N-LINKED (GLCNAC...) (POTENTIAL).
 SO SEQUENCE 95 AA; 9990 MW; 58BAF723BD7AD7F3 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 95;

Best Local Similarity 17.9%; Pred. No. 36;
 Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;
 OY 2 XXXXXXXXXXXXXXXXXXXXXXXC 29
 DB 16 LNCOTPCRADQCAENLKCRACGCAIC 43

RESULT 9
 ID RSL1_LYCES STANDARD; PRT; 96 AA.
 AC P47926;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE RSI-1 PROTEIN PRECURSOR (TRI32).
 GN RSI-1.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Solanum.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. VENT; TISSUE-ROOT;
 RX MEDLINE-94232181; PubMed-8177211;
 RA Taylor B.H., Scheuring C.F.;
 RT "A molecular marker for lateral root initiation: the RSI-1 gene of
 RT tomato (Lycopersicon esculentum Mill) is activated in early lateral
 RT root primordia.";
 RL Mol. Gen. Genet. 243:148-157(1994).
 CC -1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: EXPRESSED VERY EARLY IN LATERAL ROOT
 CC DEVELOPMENT.
 CC -1- INDUCTION: BY AUXINS.
 CC -1- PTM: SIX DISULFIDE BONDS MAY BE PRESENT.
 CC -1- SIMILARITY: BELONGS TO THE GAST1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U22188; AAA20129.1; -
 DR EMBL: U22189; AAA20130.1; -
 KW Signal.
 FT SIGNAL 1 29 POTENTIAL.
 FT CHAIN 30 96 RSI-1 PROTEIN.
 SO SEQUENCE 96 AA; 10839 MW; C17C826648148CE3 CRC64;

Query Match 88.3%; Score 68; DB 1; Length 96;
 Best Local Similarity 17.9%; Pred. No. 36;
 Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;
 OY 2 XXXXXXXXXXXXXXXXXXXXXXXC 29
 DB 36 TDCKPACTRCATSHKPKPMFFCQKC 63

RESULT 10
 ID GAST1_ARATH STANDARD; PRT; 98 AA.
 AC P46689;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GIBBERELLIN-REGULATED PROTEIN 1 PRECURSOR.
 GN GAST1
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

```

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA; TISSUE=STILLOUES;
RX MEDLINE=95244835; PubMed=7727751;
RA Herzog M., Dorne A.-M., Grellet F.;
RT "GAS1, a gibberellin-regulated gene family from Arabidopsis thaliana
  related to the tomato GAS1 gene.";
RL Plant Mol. Biol. 27:743-752(1995).
RN [2]
RP SEQUENCE OF 57-98 FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Raynal M., Grellet F., Landie M., Meyer Y., Cooke R., Delzeny M.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAY FUNCTION IN SOME HORMONAL CONTROLLED STEP OF FRUIT
  DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN FLOWER BUDS AND GREEN STILLOUES.
CC -1- LOWER LEVELS SEEN IN THE ROOTS AND ROSETTES.
CC -1- INDUCTION: BY GIBBERELLINS.
CC -1- PTM: SIX DISULFIDE BONDS MAY BE PRESENT.
CC -1- SIMILARITY: BELONGS TO THE GAS1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC
CC DR EMBL: U11766; AAB06310.1; -
CC DLR EMBL: 229940; CAA82842.1; -
CC DR MultiGene family: Signal.
CC FT SIGNAL 1 23 POTENTIAL.
CC FT CHAIN 24 98 GIBBERELLIN-REGULATED PROTEIN 1.
CC SQ SEQUENCE 98 AA: 10718 MW: 8837.550 kDa; CRC64:

Query Match 88.3%; Score 68; DB 1; Length 98;
Best Local Similarity 17.9%; Pred. No. 37;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

QY 2 XXXXXXXXXXXXXXXXXXXXXXXX 29
  ::::::::::::::::::::::::::::
DB 38 IDGSGACVACRLSRPRLCHACGTC 65

RESULT 11
GAS2_ARATH STANDARD: PRT; 99 AA.
AC P46688;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE GIBBERELLIN-REGULATED PROTEIN 2 PRECURSOR.
GN GAS2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA; TISSUE=SEED;
RX MEDLINE=95244835; PubMed=7727751;
RA Herzog M., Dorne A.-M., Grellet F.;
RT "GAS2, a gibberellin-regulated gene family from Arabidopsis thaliana
  related to the tomato GAS1 gene.";
RL Plant Mol. Biol. 27:743-752(1995).
CC -1- FUNCTION: INVOLVED IN LATE STAGES OF SEED MATURATION, OR IN EARLY
  STEPS OF GERMINATION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: SILLOUES AND DRY SEEDS.
CC -1- PTM: SIX DISULFIDE BONDS MAY BE PRESENT.
CC -1- SIMILARITY: BELONGS TO THE GAS1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC
CC DR EMBL: U11764; AAB06308.1; -
CC DLR EMBL: 229940; CAA82842.1; -
CC DR MultiGene family: Signal.
CC FT SIGNAL 1 18 POTENTIAL.
CC FT CHAIN 19 99 GIBBERELLIN-REGULATED PROTEIN 3.
CC SQ SEQUENCE 99 AA: 10704 MW: 8454.987 kDa; CRC64:

Query Match 88.3%; Score 68; DB 1; Length 99;
Best Local Similarity 17.9%; Pred. No. 37;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

```

```

CC -1- TISSUE SPECIFICITY: DRY SEEDS AND MATURING SILLOUES.
CC -1- PTM: SIX DISULFIDE BONDS MAY BE PRESENT.
CC -1- SIMILARITY: BELONGS TO THE GAS1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC
CC DR EMBL: U11765; AAB06309.1; -
CC DLR EMBL: 229940; CAA82842.1; -
CC DR MultiGene family: Signal.
CC FT SIGNAL 1 18 POTENTIAL.
CC FT CHAIN 19 99 GIBBERELLIN-REGULATED PROTEIN 2.
CC SQ SEQUENCE 99 AA: 10531 MW: 63713.986 kDa; CRC64:

Query Match 88.3%; Score 68; DB 1; Length 99;
Best Local Similarity 17.9%; Pred. No. 37;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

QY 2 XXXXXXXXXXXXXXXXXXXXXXXX 29
  ::::::::::::::::::::::::::::
DB 39 IDGGRCKDKCSKSRWKLCRLACNSCC 66

RESULT 12
GAS3_ARATH STANDARD: PRT; 99 AA.
AC P46687;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE GIBBERELLIN-REGULATED PROTEIN 3 PRECURSOR.
GN GAS3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA; TISSUE=SEED;
RX MEDLINE=95244835; PubMed=7727751;
RA Herzog M., Dorne A.-M., Grellet F.;
RT "GAS3, a gibberellin-regulated gene family from Arabidopsis thaliana
  related to the tomato GAS1 gene.";
RL Plant Mol. Biol. 27:743-752(1995).
CC -1- FUNCTION: INVOLVED IN LATE STAGES OF SEED MATURATION, OR IN EARLY
  STEPS OF GERMINATION.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: SILLOUES AND DRY SEEDS.
CC -1- PTM: SIX DISULFIDE BONDS MAY BE PRESENT.
CC -1- SIMILARITY: BELONGS TO THE GAS1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
-----
CC
CC DR EMBL: U11764; AAB06308.1; -
CC DLR EMBL: 229940; CAA82842.1; -
CC DR MultiGene family: Signal.
CC FT SIGNAL 1 18 POTENTIAL.
CC FT CHAIN 19 99 GIBBERELLIN-REGULATED PROTEIN 3.
CC SQ SEQUENCE 99 AA: 10704 MW: 8454.987 kDa; CRC64:

Query Match 88.3%; Score 68; DB 1; Length 99;
Best Local Similarity 17.9%; Pred. No. 37;
Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

```


CC -----
 DR EMBL: X63093; CAA44807.1; -.
 DR PIR: S22151; S22151.
 KW Signal.
 FT SIGNAL. 1 ? POTENTIAL.
 FT CHAIN ? 112 GASTI PROTEIN.
 SQ SEQUENCE 112 AA; 12784 MW; D369F466DF7E1C97 CRC64;

Query Match 88.3%; Score 68; DB 1; length 112;
 Best Local Similarity 17.9%; Pred. No. 40;
 Matches 5; Conservative 23; Mismatches 0; Indels 0; Gaps 0;

QY 2 XXCXXCXXXXXXXXXXCXXCXXC 29
 :|::|::|::|::|::|::|::|::|::|
 DB 52 ODCPKCTYRCSTYKPCMPFCOKKC 79

Search completed: March 1, 2001, 16:26:03
 Job time: 400 sec